MLS-C01 Q&As

AWS Certified Machine Learning - Specialty (MLS-C01)

Pass Amazon MLS-C01 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.passapply.com/aws-certified-machine-learning-specialty.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Amazon Official Exam Center

- **Instant Download** After Purchase
- **100% Money Back** Guarantee
- **365 Days** Free Update
- **800,000+** Satisfied Customers

QUESTION 1

A web-based company wants to improve its conversion rate on its landing page. Using a large historical dataset of customer visits, the company has repeatedly trained a multi-class deep learning network algorithm on Amazon SageMaker.

However, there is an overfitting problem: training data shows 90% accuracy in predictions, while test data shows 70% accuracy only.

The company needs to boost the generalization of its model before deploying it into production to maximize conversions of visits to purchases.

Which action is recommended to provide the HIGHEST accuracy model for the company\\'s test and validation data?

A. Increase the randomization of training data in the mini-batches used in training.

B. Allocate a higher proportion of the overall data to the training dataset

C. Apply L1 or L2 regularization and dropouts to the training.

D. Reduce the number of layers and units (or neurons) from the deep learning network.

Correct Answer: C

If this is a ComputerVision problem augmentation can help and we may consider A an option. However in analyzing customer historic data, there is no easy way to increase randomization in training. If you go deep into modelling and coding. When you build model with tensorflow/pytorch, most of the time the trainloader is already sampling in data in random manner (with shuffle enable). What we usually do to reduce overfitting is by adding dropout.

QUESTION 2

A manufacturer is operating a large number of factories with a complex supply chain relationship where unexpected downtime of a machine can cause production to stop at several factories. A data scientist wants to analyze sensor data from the factories to identify equipment in need of preemptive maintenance and then dispatch a service team to prevent unplanned downtime. The sensor readings from a single machine can include up to 200 data points including temperatures, voltages, vibrations, RPMs, and pressure readings.

To collect this sensor data, the manufacturer deployed Wi-Fi and LANs across the factories. Even though many factory locations do not have reliable or high-speed internet connectivity, the manufacturer would like to maintain near-real-time inference capabilities.

Which deployment architecture for the model will address these business requirements?

A. Deploy the model in Amazon SageMaker. Run sensor data through this model to predict which machines need maintenance.

B. Deploy the model on AWS IoT Greengrass in each factory. Run sensor data through this model to infer which machines need maintenance.

C. Deploy the model to an Amazon SageMaker batch transformation job. Generate inferences in a daily batch report to identify machines that need maintenance.

D. Deploy the model in Amazon SageMaker and use an IoT rule to write data to an Amazon DynamoDB table.

Consume a DynamoDB stream from the table with an AWS Lambda function to invoke the endpoint.

Correct Answer: B

https://aws.amazon.com/blogs/iot/using-aws-iot-for-predictive-maintenance/ https://aws.amazon.com/blogs/iot/industrial-iot-from-condition-based-monitoring-to-predictive-quality-to-digitize-your-factory-with-aws-iot-services/

QUESTION 3

A machine learning specialist is running an Amazon SageMaker endpoint using the built-in object detection algorithm on a P3 instance for real-time predictions in a company\\'s production application. When evaluating the model\\'s resource utilization, the specialist notices that the model is using only a fraction of the GPU.

Which architecture changes would ensure that provisioned resources are being utilized effectively?

A. Redeploy the model as a batch transform job on an M5 instance.

B. Redeploy the model on an M5 instance. Attach Amazon Elastic Inference to the instance.

C. Redeploy the model on a P3dn instance.

D. Deploy the model onto an Amazon Elastic Container Service (Amazon ECS) cluster using a P3 instance.

Correct Answer: B

https://aws.amazon.com/machine-learning/elastic-inference/

QUESTION 4

A data scientist at a financial services company used Amazon SageMaker to train and deploy a model that predicts loan defaults. The model analyzes new loan applications and predicts the risk of loan default. To train the model, the data scientist manually extracted loan data from a database. The data scientist performed the model training and deployment steps in a Jupyter notebook that is hosted on SageMaker Studio notebooks. The model\\'s prediction accuracy is decreasing over time.

Which combination of steps is the MOST operationally efficient way for the data scientist to maintain the model\\'s accuracy? (Choose two.)

A. Use SageMaker Pipelines to create an automated workflow that extracts fresh data, trains the model, and deploys a new version of the model.

B. Configure SageMaker Model Monitor with an accuracy threshold to check for model drift. Initiate an Amazon CloudWatch alarm when the threshold is exceeded. Connect the workflow in SageMaker Pipelines with the CloudWatch alarm to automatically initiate retraining.

C. Store the model predictions in Amazon S3. Create a daily SageMaker Processing job that reads the predictions from Amazon S3, checks for changes in model prediction accuracy, and sends an email notification if a significant change is detected.

D. Rerun the steps in the Jupyter notebook that is hosted on SageMaker Studio notebooks to retrain the model and redeploy a new version of the model.

E. Export the training and deployment code from the SageMaker Studio notebooks into a Python script. Package the

script into an Amazon Elastic Container Service (Amazon ECS) task that an AWS Lambda function can initiate.

Correct Answer: AB

QUESTION 5

A Machine Learning Specialist is working with a large cybersecurily company that manages security events in real time for companies around the world The cybersecurity company wants to design a solution that will allow it to use machine learning to score malicious events as anomalies on the data as it is being ingested The company also wants be able to save the results in its data lake for later processing and analysis

What is the MOST efficient way to accomplish these tasks\\'?

A. Ingest the data using Amazon Kinesis Data Firehose, and use Amazon Kinesis Data Analytics Random Cut Forest (RCF) for anomaly detection Then use Kinesis Data Firehose to stream the results to Amazon S3

B. Ingest the data into Apache Spark Streaming using Amazon EMR. and use Spark MLlib with k-means to perform anomaly detection Then store the results in an Apache Hadoop Distributed File System (HDFS) using Amazon EMR with a replication factor of three as the data lake

C. Ingest the data and store it in Amazon S3 Use AWS Batch along with the AWS Deep Learning AMIs to train a k-means model using TensorFlow on the data in Amazon S3.

D. Ingest the data and store it in Amazon S3. Have an AWS Glue job that is triggered on demand transform the new data Then use the built-in Random Cut Forest (RCF) model within Amazon SageMaker to detect anomalies in the data

Correct Answer: A

https://aws.amazon.com/tw/blogs/machine-learning/use-the-built-in-amazon-sagemaker-random-cut-forest-algorithm-for-anomaly-detection/